# Reference/Citation Linking and CrossRef

June 21, 2000


**Ed Pentz**

**Executive Director**

**CrossRef**

“Although I’ve never met anyone at a record company who ‘believed in the Internet,’ they’re all trying to cover their asses by securing everyone’s digital rights”

*Courtney Love (Salon.com)*

Publishers may or may not believe in the Internet, but they are all covering themselves by linking, especially to full text articles

- **Reference, n.**
- **5. a.** A direction to a book, passage, etc., where certain information may be found; an indication of the author, work, page, etc., to be looked at or consulted.
- BRINSLEY *Lud. Lit.* xiii. (1627) 188 If they had but only bookes of References, it would be exceeding profitable.

*Oxford English Dictionary*

# Reference Linking

- Gives end users access to logically related articles in one or two clicks
- Allows body of research literature to be seen as logically associated
- Researchers want and expect everything to link

No Links = Decreased Value

# Why CrossRef?

- Market pressure to link
- Current Linking
  - Closed Links - only within journal systems
  - Links between large secondaries and large primaries - small publishers left out, primaries not linking to each other
- Linking Agreements
  - Bilateral agreements
  - Define terms of linking/“no surprises”

# Why CrossRef?

- 2-party agreements not scalable
  - too many publishers (primary and secondary)
  - ISI has linking agreements with 15 publishers
- Result - Publishers cooperating!
- CrossRef makes broad-based linking manageable

# Reference Linking Infrastructure

- How is a link to an article created from a standard citation?
  - Algorithmic Link - bibliographic metadata
  - Article Identifier w/ resolution system
- To create algorithmic links -
  - a publisher needs to know the linking scheme of every cited publisher and when it changes
  - publisher needs to know if the article is online - massive data collection problem

# Article Identifier

- Digital Object Identifier (DOI)
  - Handle System - CNRI
  - unique, persistent, soon to be NISO standard
  - DOI System routes a DOI to a URL registered by publisher
- With a DOI -
  - linker doesn’t need to know publishers’ linking systems or if they change
  - if a publisher changes their URLs, they only need to be updated in one place
  - article is guaranteed to be online (only articles online get DOIs)

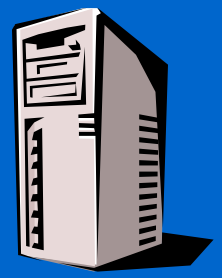
# DOI

- DOI Problem
  - how do you know what the DOI is for an article?
  - CrossRef provides a database of DOIs and a DOI Lookup service (telephone book and directory assistance)

# The CrossRef Process

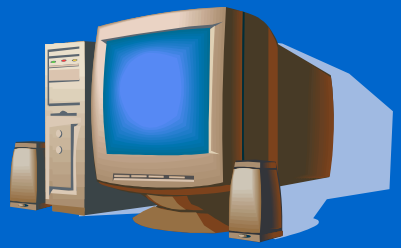
- Publishers deposit article metadata (including DOIs) in CrossRef database
- Publishers submit references from articles to CrossRef to lookup DOIs
- Publishers use DOIs to create reference links in their online journals
- User clicks on reference links and goes to other Publisher's site

# WORKFLOW FOR REFERENCE LINKING



Publisher Production System

3. Journal production systems export articles with DOIs

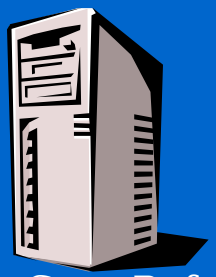


Publisher Online Journal System

2. DOI Lookup: a Journal production system requests a DOI based on citation metadata

1. Journal production systems export metadata to MDDB as new articles are published

4. Users may request and retrieve articles via DOI resolution system



CrossRef MDDB

User

# CrossRef

- Not-for-profit, membership organization
  - Members: publishers of original scholarly material
  - Users: members and any organization creating links to full text articles
  - Purpose 1: Enable links from primary article references to the cited articles at other publishers' sites
  - Purpose 2: Enable links to full text articles from secondary services

# Organization

- Publishers International Linking Association (PILA) - incorporated February 2000
- Board of Directors
  - AAAS (Science),Academic Press (Harcourt), AIP, ACM, Blackwell Science, Elsevier Science, IEEE, Kluwer, Nature, OUP, Springer, Wiley

# Business Rules/Governance

- CrossRef Membership
  - **Provide full bibliographic citation for incoming DOI links**
  - many publishers will give free abstracts
  - information on acquiring article (pay online, document delivery, subscription)
  - Access to abstracts and full text controlled by publishers
- CrossRef guarantees links

# Current Status

- Reference links are live!!
- 1100 journals have links (Elsevier, Academic Press, Wiley, Springer, Blackwell Science)
- Metadata deposited for 1.3 Million articles from 2700 journals
- 3 million articles by end of year
- 0.5-1 million new articles per year

- AAAS (Science)
- Academic Press (Harcourt)
- ALPSP
- American Chemical Society
- American Geophysical Union
- American Institute of Physics
- American Mathematical Society
- American Physical Society
- American Psychological Assoc
- ACM
- Annual Reviews
- Blackwell Science/Blackwell Publishers
- British Medical Journal
- CAB International
- Cambridge University Press
- Elsevier Science
- IEE
- IEEE
- Institute of Physics
- Kluwer Academic
- Marcel Dekker
- Mary Ann Liebert
- Nature
- Oxford University Press
- Portland Press
- The Royal Society
- Royal Society of Chemistry
- Springer
- Taylor & Francis
- Thieme Verlag
- University of Chicago Press
- John Wiley & Sons
- World Scientific

crossref

# CrossRef Fees

- Member Fee
- Annual Administrative Fee
- Deposit Fee
- Lookup Fee
- Principles
  - cost recovery
  - flexibility
  - no charges to end users to follow links

# Fees

- **Primary Publisher Annual Member Fee**
  - 1 title, max 500 articles per year *$200*
  - 2-5 titles, max 2,500 articles per yr *$500*
  - 6-20 titles, max 10,000 articles per yr *$750*
  - 21-100 titles, max 50,000 articles per yr *$1,000*
  - >100 titles or >50,000 articles per yr *$2,000*

- **Non-Member Annual Administrative Fee**
  - Secondary, database <100,000 records/yr *$2,000*
  - Secondary, database >100,000 records/yr *$5,000*
  - Agents *$5,000*
  - Non-resellers (e.g., libraries) *$300*

# Fees

- **Deposit Fee for Primary Material**
- Current Year -- full text $0.60
- Back File (deposit after 10/1/2000) $0.10
- Back File (prior to 10/1/2000) $0.05

# Fees

- **Member Retrieval Fee** (for successful match)
- Per DOI matched $0.05

- **Non-Member Retrieval Fee** (for successful match)
- Per DOI Matched $0.10

# Future Developments

- Multiple Resolution
  - one DOI = one URL is starting point
  - multiple locations and multiple files
- Appropriate Copy Issue
- Collaboration
  - DLF/IDF/CNRI/CrossRef, libraries, publishers
- Interoperability
  - Metadata is “system neutral”

# Contact Details

- www.crossref.org
- Ed Pentz - epentz@crossref.org

# Metadata Deposit

- Publishers submit article metadata
  - Journal Title, ISSN/Coden,Volume, Issue, First Page, First Author, Year - mandatory
  - Article Title is optional
- XML-based DTD (DOI-X)
  - Formal grammar
  - Supplementary “data rules”
- Journal articles today
  - conference proceedings and more to follow

# DOI Lookup

- Publishers submit references to Reference Resolver (RR)
  - References tagged in SGML/XML
  - Minimum Data
    - Abbreviated title, volume, first page - possibly first author and year
  - Send Batches of References
  - RR - intelligent “link agent”